Ali Garjani

garjania.github.io

ali.garjani@epfl.ch
Ali Garjani
garjania
ali-garjani

Education

2021-2024 **EPFL (École Polytechnique Fédérale de Lausanne)**, *Switzerland*,
MSc. Computational Science and Engineering | GPA: **5.59/6**,
Master Thesis: Investigating Tokenization Techniques for Improving Multimodal Foundation Models

2017-2021 **Sharif University of Technology**, *Iran*,
BCs. Computer Engineering | GPA: **19.08/20 (Top 10%)** ,
Bachelor Thesis: Neural Distributed Image Compression using Common Information

Research Interests

- **Multimodal Learning:** Enhancing vision-based models by integrating multimodal data to improve scene perception and understanding.
- **Learning from Dynamic Data:** Utilizing temporal and spatial information from videos, multi-view, and 3D data to improve the robustness and the performance of models across diverse tasks.
- **World Modeling:** Training models to understand the physcis of the world through future-state predcition.

Publications & Preprints

• **Multimodality as Supervision: Self-Supervised Specialization to the Test Environment via Multimodality**,
Kunal Pratap Singh, **Ali Garjani***, Muhammad Uzair Khattak, Rishubh Singh, Andrei Atanov, Oğuzhan Fatih Kar, Amir Zamir*,
– arxiv 2025, [Website]

• **How Well Does GPT-4o Understand Vision? Evaluating Multimodal Foundation Models on Standard Computer Vision Tasks**,
*Rahul Ramachandran, **Ali Garjani**, Andrei Atanov, Oğuzhan Fatih Kar, Amir Zamir*,
– arxiv 2025, [Website][Paper][GitHub]

• **4M-21: An Any-to-Any Vision Model for Tens of Tasks and Modalities**,
Roman Bachmann, Oğuzhan Fatih Kar*, David Mizrahi*, **Ali Garjani**, Mingfei Gao, David Griffiths, Jiaming Hu, Afshin Dehghan, Amir Zamir*,
– Neural Information Processing Systems (**Neurips**), 2024, [Website][Paper][GitHub][Demo]

• **Neural Distributed Image Compression with Cross-Attention Feature Alignment**,
Nitish Mital, Ezgi Özyılkan*, **Ali Garjani***, Deniz Gunduz*,
– Winter Conference on Applications of Computer Vision (**WACV**), 2023, [Paper][GitHub]

• **Neural Distributed Image Compression using Common Information**,
Nitish Mital, Ezgi Özyılkan*, **Ali Garjani***, Deniz Gunduz*,
– Data Compression Conference (**DCC**), 2022, [Paper][GitHub]

• **Forecasting Influenza Hemagglutinin Mutations Through the Lens of Anomaly Detection**,
***Ali Garjani**, Atoosa Malemir Chegini, Mohammadreza Salehi, Alireza Tabibzadeh, Parastoo Yousefi, Mohammad Hossein Razizadeh, Moein Esghaei, Maryam Esghaei, Mohammad Hossein Rohban*,
– **Nature - Scientific Reports Journal**, 2023, [Paper][GitHub]

• **Importance of OCT-derived Biomarkers for the Recurrence of Central Serous Chorioretinopathy using Statistics and Predictive Modelling**,
*Emilien Seiler, Léon Delachaux, Jennifer Cattaneo, **Ali Garjani**, Alexia Duriez, Thibaud Martin, Jérémy Baffou, Sepehr Mousavi, Ilenia Meloni, Ciara Bergin, Mattia Tomasoni, Chiara M Eandi*,
– **Nature - Scientific Reports Journal**, 2024

* Equal contribution, randomized author order

Research Experience

2024 Mar - Nov **Research Intern at Visual Intelligence and Learning Lab**, *EPFL*,
Advised by Prof. Amir Zamir

- **Multimodality as Supervision:** Leading a project that utilizes self-supervised multimodal learning to specialize the model for a specific physical test space, enabling it to effectively perform unseen tasks in that space.
- **Understanding the Visual Capabilities of MLLMs**: Evaluating of multimodal large language models such as ChatGPT, Gemini, and Claude on well-known vision tasks and benchmarks.

2023 Feb - Dec
Research Lab Assistant at Visual Intelligence and Learning Lab, *EPFL*,
Advised by Prof. Amir Zamir
- **Scaling Multimodal Foundation Models:** Scaling a multimodal multitask model to handle 21 diverse modalities, using discrete tokenization to unify representations and enhance training scalability.

2020-2021
Research Intern at Information Processing and Communications Lab, *Imperial College London*,
Advised by Prof. Deniz Gunduz
- **Neural Image Compression:** Designed an autoencoder architecture and formulated a loss function for efficient distributed neural image compression using side information.

Work Experience

2025
Research Intern and ML Engineer at Duranta, *Seattle*,
Building robust segmentation models from sparse signals.

2023-2024
Research Lab Assistant and Intern at EPFL, *Lausanne*,
Training and experimenting with multimodal foundation models.

2022
Data Science Intern at Jules Gonin Eye Hospital, *Lausanne*,
Trained a model to predict a disease recurrence in patients using their OCT scans.

2021
Software Engineer at Wize Analytics, *Tehran*,
Designed and implemented an automatic data engineering pipeline.

Teaching Experience

EPFL
AI Product Management, Image Processing II
Participated as a mentor in exercise classes

Sharif University
Artificial Intelligence, Linear Algebra (Practical Head)
Designed problems for exams and assignments, and mentored exercise classes

Specialized Skills

Machine Learning
Multimodal Learning, Self-Supervised Learning, Diffusion Models, Tokenization

Theory & Math
Numerical Partial Differential Equations, Complex Dynamical Systems Analysis, Information Theory

Technical Skills

AI Toolkits
AI2-Thor, Detectron2, Open-MMLab

Library
Pytorch, Tensorflow, Transformers, Ray Serve, Selenium

Software
Blender, Unreal Engine, Adobe

Frameworks
Git, Docker, Django, Android Studio, React Native

Programming Languegs
Python, C++, C, Java, Swift, R, Javascript, Matlab

Honors and Awards

- Recepient of the EDIC Fellowship (For the first year of PhD studies).
- Ranked 126th in nation-wide Iranian Universities Entrance Examination among more than 148000 participants (2017).

Academic Service

Presentations
Robot Learning Lab, University of Freiburg, *Presented the works 4M and Multimodality as Supervision*

Presentations
EPFL AI Day, *Presented 4M and ViPer to the public*

Reviewer
CVPR 2025, ICLR 2025, AAAI 2025, CVPR 2024, DCC 2023

Scientific Staff
Sharif FieldIn, DataDays, Webelopers, AI Challenge

Languages & Hobbies

Language
Persian (Native), English (Advanced), Azari (Fluent)

Sports
Bouldering, Tennis, Football, Volleyball

Art
Creative Coding